

《教育測驗與統計概要》

一、某教授想瞭解班上100位學生是否精熟，他請了A、B兩位評分員協助評定學生成就，評定結果如下：

	A評分員	
B評分員	精熟	不精熟
精熟	40	20
不精熟	10	30

(一)有那些方法可以瞭解此測驗的信度？(10分)

(二)請使用這些方法，計算此測驗的信度。(14分)

試題評析	本題成就測驗信度的探討，由於以學生精熟與否評定，屬效標參照測驗，通常是考生比較忽略準備的考試範圍，雖然不難，然解題上會稍費工夫。但是，本班考生因平日上課有所講解與練習，應較能掌握分數。
考點命中	《教育測驗與統計講義第三回》，傅立葉編撰，第14章，頁14。

答：

(一)依題意敘述，如果該教授要了解所施測成就測驗的信度，可採用的方法有：

- 1.由於該教授對於學生學習成就的評量目標為「是否精熟」，此為二分名目尺度的概念，而且具備效標參照測驗的屬性，因此，「百分比一致性」(PA)索求算之兩位評分員的評定一致性可以被使用。
- 2.另外，科恆(Cohen)的Kappa係數也可利用。

(二)以前述方法，可以計算此測驗的信度如下：

$$1. P(A) = \frac{a}{N} + \frac{d}{N} = \frac{40 + 30}{40 + 20 + 10 + 30} = 0.7$$

$$2. P(E) = \left(\frac{a+c}{N}\right)\left(\frac{a+b}{N}\right) + \left(\frac{b+d}{N}\right)\left(\frac{c+d}{N}\right) = \left(\frac{50}{100}\right)\left(\frac{60}{100}\right) + \left(\frac{50}{100}\right)\left(\frac{40}{100}\right) = 0.5$$

$$K = \frac{P(A) - P(E)}{1 - P(E)} = \frac{0.7 - 0.5}{1 - 0.5} = 0.4$$

二、(一)如果想要計算非對即錯試題的難度，可使用那些指數？請說明這些指數的意義。(16分)

(二)如果想要計算非對即錯試題的鑑別度，可使用那些指數？請說明這些指數的意義。(10分)

試題評析	本題的兩個子題針對非對即錯試題的難度與鑑別度分析指數命題，屬於較容易且所有考生都能夠容易掌握分數的題目之一。得分高低的關鍵在於，考生能否抽絲剝繭，一層一層的詳細敘述作答。
考點命中	《教育測驗與統計講義第三回》，傅立葉編撰，第18章，頁32~33。

答：

(一)對於非對即錯試題的難度計算、解釋與探討，可以使用的難度指數與其意義，分別說明如下：

1.通過百分比指數：

a.計算全體受試者答對每題的人數百分比，稱為難度指數。可以計算如下：

$$P=R/N$$

優點：簡單易行。

缺點：無法顯示試題區別功能，到底是高分組或低分組答對較多。

b.先分別計算高分組與低分組在每一道題答對人數與百分比。然後計算：

$$P=(PH+PL)/2$$

優點：

- (1)答對百分比表示難度，其數值愈大，題目愈簡單。
- (2)使用次序量尺，可以指出題目的等級順序或相對難度。

缺點：

- (1)無法指出各難度間差異大小。
- (2)單位不相等，即 $P_1-P_2 \neq P_2-P_3$

$$2.\Delta=13+4z$$

數字愈大，代表題目愈困難；因屬於等距量尺，不同試題之間的難度可直接做比較。

(二)鑑別力分析的目的，在於確定題目是否具有區分受試者能力程度高下的作用。依分析時所依據的標準可有內部一致性分析與外在效度分析兩種方式。計算非對即錯試題的鑑別度，可以使用的指數與其意義，分別說明如下：

1.內部一致性分析

旨在瞭解各個試題的功能是否和整個測驗的功能相符合一致。公式如下：

$$D=PH-PL$$

鑑別度指數的取值域： $-1 \leq D \leq 1$

- a. $D=0$ ，無鑑別度。不是試題太簡單，高分組與低分組學生全部答對($PH-PL=0$)。就是試題太困難，高分組與低分組學生全部答錯($PH-PL=0$)。
- b. $D=+1$ ，高分組學生全部答對，低分組學生全部答錯。 $(PH=1, PL=0)$ 。
- c. $D=-1$ ，低分組學生全部答對，高分組學生全部答錯。 $(PH=0, PL=1)$ 。

2.外在效度分析

旨在檢驗題目是否具有預定的某種鑑別作用。根據學生在某試題的反應對或錯，求算其與測驗總分之相關係數，可用來表示試題鑑別度指標。依測驗結果的人為二分或自然二分，常用的方法有：點二系列相關與二系列相關等。

三、(一)請舉例說明正偏態 (positively skewed) 的意義。(5分)

(二)請舉例說明矩形分配 (rectangular distribution) 的意義。(5分)

(三)請舉例說明二項式分配 (binomial distribution) 的意義。(4分)

(四)請舉例說明二項式分配母群 (population) 的平均數與標準差分別為何？(6分)

(五)變異數分析的結果若達到統計的顯著性，可接著做事後比較，事後比較包含簡單比較與複雜比較，請解釋何謂簡單比較？(5分)

試題評析	本題考資料分配的正偏，以及變異數分析的事後比較，都是教育統計的重要觀念與方法，考生不會有太大疑問。出乎意料地，矩形分配與二項分配這兩個罕見在公職考試的連續型與間斷型機率分配竟佔了15分。沒有相關完整專業課程訓練考生會在本題徹底被擊倒。
考點命中	1.《教育測驗與統計講義第一回》，傅立葉編撰，第2章，頁6。 2.《教育測驗與統計講義第二回》，傅立葉編撰，第13章，頁48。

答：

(一)正偏態意旨資料的大小分配，呈現數據小者遠多於數據大者的現象；若依三個集中量數的大小關係來說，則有平均數最大，眾數最小的情形；當以圖形來看，則有低而緩長的右尾。例如：100位小學資源班學生的數學測驗成績，多數偏低分，極少數高分所呈現的即為正偏態。

(二)矩形分配可以先從均勻分配談起，均勻分配指的是在X的值域範圍內，其範圍的機率皆相等，但在X的值域範圍外，則機率為零。它有兩個參數，分別代表值域的起點與終點，用符號表示成 $X \sim U(a, b)$ 。因為值域範圍內的機率皆相同，故又稱矩形分配。

以等公車為例，令X為等公車的時分，公車平均每十分鐘來一班，因此 $X \sim U(0, 10)$ ，則等待公車時分X的機率為 $f(X)=1/(10-0)=1/10, 0 < X < 10$ 。

(三)二項式分配源自伯努利試驗(Bernoulli Trial)，其特色為：每一次試驗皆僅有兩種可能結果，不是成功(S)，就是失敗(F)。且成功機率固定為 $P(S)=p$ ，失敗機率固定為 $P(F)=1-p$ 。加上，每一次試驗之間互為獨立。進行n次的伯努利試驗，稱為二項實驗，若隨機變數X為n次試行實驗成功的次數，X的機率分配稱為二項機

率分配。例如，求職者的性向測驗是否及格的決定，單一求職者能夠順利及格或未能通過的機率探討適用此一分配。

(四)二項式分配母群的平均數與標準差分別為：

X的期望值為 np ；變異數為 $np(1-p)$ ，因此標準差為 $np(1-p)$ 的平方根。

例如：現今大學生中成功升學研究所的機率已知為0.65，則在已經隨機抽出的100位大學生中，平均而言有

$100(0.65)=65$ 位同學可以考上研究所，而考上人數的標準差為 $\sqrt{100(0.65)(0.35)} = 4.769$ 人

(五)變異數分析達統計顯著性時，所接續進行的簡單事後比較指的是，事後一一比較所有兩兩類別之間平均數的差異，然後讓我們知道是哪兩個類別間平均數的差異對於做ANOVA測定時得到之F值的貢獻最大。由於做兩兩類別間之比較時，我們犯TYPE I錯誤的機率增加，因此事後分析法會以更嚴格的標準來判定哪兩各類別之平均數間的差異達到顯著水準。

例如：三個抽自台灣北、中、南三區的小學生班級，當其價值觀量表平均分數不全等的假設考驗達顯著，則後續進行兩兩區域平均價值觀是否相等的事後比較。

四、某研究者想瞭解A、B兩班的變異程度是否不同，而A班共有13人，A班的標準差是20，B班共有20人，標準差是10。請使用假設考驗的步驟，考驗A、B兩班的變異程度不同的假設，並解釋研究結果。(F. $_{975}(12, 19) = 2.72$) (25分)

試題評析	本題考的是兩個母體變異數相等與否的顯著性考驗。由於題目提供臨界值的數據，考生減少查表時可能犯錯的風險，因此題目分數是今年四題中最容易取得者。
考點命中	《教育測驗與統計講義第二回》，傅立葉編撰，第11章，頁22。

答：

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

$$F = 20^2 / 10^2 = 4, 4 > 2.72,$$

拒絕虛無假設，A、B兩班學生變異程度顯著不同。

【版權所有，重製必究！】